From Expanded Cinema to Extended Reality: How AI Can Expand and Extend Cinematic Experiences

Junrong Song
Computational Media and Arts, Hong Kong University of Science and Technology (Guangzhou)
jsong799@connect.hkust-gz.edu.cn

Bingyuan Wang
Computational Media and Arts, Hong Kong University of Science and Technology (Guangzhou)
bwang667@connect.hkust-gz.edu.cn

Zeyu Wang
Computational Media and Arts, Hong Kong University of Science and Technology (Guangzhou)
zeyuwang@ust.hk

David Kei-man Yip*
Computational Media and Arts, Hong Kong University of Science and Technology (Guangzhou)
daveyip@ust.hk

Figure 1: XR film shooting set with AI-generated environment in campus.

ABSTRACT

This paper explores the concept of expanded cinema and its relationship to extended reality (XR), focusing on the potential of artificial intelligence (AI) to expand and extend expressive possibilities. Expanded cinema refers to experimental film and multimedia art forms that challenge the conventions of traditional cinema by creating immersive and interactive experiences for audiences. XR, on the other hand, blurs the line between physical and virtual reality, offering immersive storytelling experiences. Both expanded cinema and XR aim to push the boundaries of traditional norms and create immersive experiences through the integration of technology, interactivity, and cross-sensory elements. The paper emphasizes the role of AI in optimizing 3D scene creation for XR and enhancing the overall experience through a case study. It also presents several AI-based techniques, such as generative models and AI-assisted rendering, that facilitate efficient and effective 3D content creation. Additionally, it explores the use of AI plugins in 3D modeling software and the generation of 3D models and textures from 2D images using techniques like GANs and VAEs. The incorporation of AI to extend and expand opens up new possibilities for immersive experiences in the future.

CCS CONCEPTS

• **Applied computing**; • **Arts and humanities**; • **Fine arts**;

KEYWORDS

Extended reality, expanded cinema, artificial intelligence.

ACM Reference Format:
Junrong Song, Bingyuan Wang, Zeyu Wang, and David Kei-man Yip. 2023. From Expanded Cinema to Extended Reality: How AI Can Expand and Extend Cinematic Experiences. In *Proceedings of The 16th International Symposium on Visual Information Communication and Interaction (VINCI 2023).* ACM, New York, NY, USA, 6 pages. https://doi.org/XXXXXXX.XXXXXXX

1 INTRODUCTION

In audio-visual media, traditional storytelling methods typically seek to elicit cognitive and affective reactions by primarily engaging the senses of sight and sound. Over the past century, cinema, as a dominant form of this media, has evolved from the hand crank projector, the first 3D glasses, and the cinemascope of the 1950s, the iMax screens, to the present emerging AIGC. Throughout the evolution of film, some experimental filmmakers have attempted to push the boundaries of traditional cinema to advance film-making. The term "expanded cinema" [55] describes experimental film and multimedia art forms that often involve multiple projectors, live performances, and other multimedia installations to create an immersive and interactive experience for audiences. American filmmaker Stan Vanderbeek coined the term in the mid-1960s to describe a

*Corresponding author

VINCI 2023, September 22–24, 2023, Guangzhou, China

ACM ISBN 978-1-4503-XXXX-X/18/06…$15.00
https://doi.org/XXXXXXX.XXXXXXX

movement of artists and filmmakers who explored new ways of experiencing film. Expanded cinema represents a departure from the passive concept of traditional single-screen film experiences; it is an experimental form of film-making that expands and challenges the boundaries of the medium, surpassing the limits of artistic vision and viewing angles.

Gene Youngblood's book has early records of using computers to create expanded cinema. One of the earliest examples was "Labyrinthe" [63] for Expo '67 at Montreal, a multi-screen presentation that integrated different images, sometimes in contrast, sometimes complementing each other. Youngblood envisioned that live simulation using advanced computers could achieve real-time interaction and immersion and believed it would coexist with traditional cinema as a new cybernetic art form.

Today's expanded cinema utilizes electronic and digital technologies to simulate extended space, time, and reality [63]. While expanded cinema, extended reality, and virtual production differ in some concepts and technical means, they share similarities in creating immersive experiences. Compared to traditional cinema, expanded cinema tried to create interactive experiences for audiences. It attempts to create illusions of reality within a space by expanding film into a spatial medium. The multiple-screen display and players connected to a programmed video mixer are comparable to today's multiple LED n-display concept. Through a powerful real-time media server, they can connect hardware like motion capture, volumetric capture, and photogrammetry into the production pipeline, providing countless creative possibilities in an immersive virtual environment — the so-called extended reality. And with the potential of using artificial intelligence to assist in expanding and extending the expressive possibilities of filmmaking.

2 EXPANDED CINEMA AS EXPANDED REALITY

The art movement of the 1960s sought to expand the boundaries of cinema beyond the screen. For instance, Stan Brakhage's films "Window Water Baby Moving" (1959) and "Mothlight" (1963) [16] were notable attempts to explore the aesthetic potential of light, color, and motion. Similarly, Michael Snow's film "Wavelength" (1967) [6] employed a series of long shots to investigate the interplay between space and time within a house, providing viewers with a multi-sensory experience of these dynamic changes.

Valie Export's lecture on "Expanded Cinema" at "The Essential Frame – Austrian Independent Film 1955–2003" states that expanded cinema is the simulation of space and time. She also refers to expanded cinema as expanded reality [32]. A separate article explores two distinct forms of expanded cinema. The first revolutionizes by altering the conditions of projection, while the second employs multiple screens and projectors to create a unique viewing experience [4]. Thus, "Expanded cinema" and "Expanded reality" are closely linked concepts referring to the use of technology to create new kinds of reality.

The book "Cinema Expanded: Avant-Garde film in the Age of Intermedia" [51, 56] states that expanded cinema has evolved since the mid-1960s. Since then, expanded cinema has taken various forms, like moving image installation, multi-screen films, live cinematic performances, shadow plays, computer-generated images, and video art. Through experimental techniques involving light, space, motion, and interactivity, expanded cinema generates new kinds of cinematic reality and prompts contemplation on the nature of film itself [63].

It can be argued that expanded cinema has opened up new avenues for developing an expanded reality. For instance, John Baldessari's 'Electronic Labyrinth' [25] used multiple projectors and lights to create interactive multimedia installations - blurring the boundary between artwork and viewer. It was a successful early example of interactive cinema too. It highlighted the potential of interactivity, space, and light - which would become crucial for expanded reality technologies. Pioneers like Anthony McCall and French expanded cinema [56] also explored how projections, light and sound, and other art forms could simulate or distort reality perception and feed into expanded reality technologies.

Expanded cinema has been expanding through technical means as an avant-garde art form, with its significance in developing a cinematic language that further expands human expressive abilities and consciousness. Pioneers in expanded cinema have been exploring these aspects for decades [50].

3 FROM EXPANDED REALITY TO EXTENDED REALITY (XR)

The latest cutting-edge media technology that blurs the line between physical and virtual reality is called Extended Reality (XR). From the 1960s to the 2000s, certain "extended reality" artworks attempted to broaden human senses and cognitive abilities by creating surreal experiences using lenses, slides, and other experimental techniques [35]. Compared to contemporary XR technologies, these extended reality techniques and experiences may have been primitive and rudimentary, but they laid the groundwork for later immersive technologies. Such as "The Tunnel under the Atlantic" created by artist Maurice Benayoun in the 1990s, showcasing early endeavors to construct digital environments [37].

Today's XR technology has also enabled immersive and expanded forms of storytelling [61], which enhance the concept of expanded cinema by allowing it to return to its original cinematic function with more ease and possibilities, allowing for the expansion of narrative and emotional experiences. Unlike traditional "extended reality" forms, XR allows the audience to further actively participate, control and influence the experience instead of only experiencing it on multiple screens. Filmmakers can use XR to create 360-degree environments that viewers can explore as the story unfolds, and visual effects can be enhanced to create a more immersive and realistic experience [11]. XR can help on 3D virtual sets, using green screens or LED video walls for live compositing with moving cameras while offering new storytelling methods. Both attempt to immerse the audience with multiple views and perspectives.

The other common feature of them is their incorporative and integrative nature. Both achieve immersive experiences by integrating different technologies. Expanded cinema integrates various media, like video displays and computers. Similarly, XR technology integrates LED display systems, studio camera tracking, lighting, motion capture, and powerful 3D tools like Unreal Engine to create more realistic and interactive virtual experiences [14]. Both

Figure 2: Showcases 3D virtual sets displayed on an LED video wall and used for live compositing with 4K moving cameras.

expanded Cinema and XR constantly expand, aiming to create experiences beyond traditional norms. And at the same time, both seek to foster better immersive experiences.

And with the development of Artificial Intelligence (AI), its role in expanded cinema and XR is becoming increasingly important. For instance, AI can help create personalized scenes and surroundings for the audience by adapting the content to their preferences and interests. As such, it has the potential to revolutionize the way we experience and interact with art, leading to more immersive and personalized experiences.

4 AI ASSISTS IN CREATING EXTENDED REALITY AND EXPANDED CINEMA TODAY

Current AI technology presents a multitude of opportunities to enhance the creation of extended cinema and extended reality. Specifically, our focus is on the challenge of 3D scene creation, a crucial aspect of XR production. Through our research, we have shown that AI can optimize this process, resulting in more efficient and realistic 3D scenes. We also provide an overview of the latest theories, models, and tools that can be utilized to tackle related problems in this area. Additionally, we highlight existing practices and present a case study conducted by our team, illustrating the potential benefits of incorporating AI into various forms of artwork and how this integration can propel AI's involvement in XR forward.

4.1 Why AI is Needed in XR

Extended reality (XR) technologies offer exciting possibilities for creating immersive and engaging experiences for users. However, creating high-quality and realistic 3D scenes for XR applications can be a time-consuming and complex task. On the other hand, there are many low-level, repetitive operations involved in the iterative creation and modification process, which can be optimized by incorporating computational technologies. This is where AI can prove invaluable, offering a more efficient and intuitive solution. By leveraging AI techniques, such as generative models and AI-based rendering, artists and designers can streamline their workflows, automate certain tasks, and achieve impressive results with greater ease and speed. AI-powered tools and platforms offer a more convenient and intuitive approach to 3D scene creation, empowering creators to focus on their artistic vision and elevating the overall XR experience for users.

Conventional AI techniques, such as traditional machine learning algorithms (e.g., decision trees, naive Bayes, *k*-means [33]) and deep learning algorithms (mainly based on neural networks, including CNN, RNN, Transformers [47]), have been widely used in 2D image and video processing tasks. However, these techniques are not directly applicable to 3D environments. The main challenge is that traditional AI methods lack an understanding of the three-dimensional representation of the environment, and thus cannot easily create or manipulate 3D objects and scenes. However, advanced AI techniques, such as neural rendering fields (NeRF) [41], possess such capabilities. Currently, computer vision algorithms can be trained to recognize and understand the 3D geometry of a scene, enabling more accurate and efficient 3D reconstruction. Generative models, such as Generative Adversarial Networks (GANs) [17] and diffusion models [22], can be trained to generate new and diverse 3D content, helping to expand the creative possibilities for XR applications.

Current XR technologies benefit greatly from the capabilities of AI in 3D scene creation. Recent advancements in AI have introduced innovative approaches for text-to-3D conversion, object generation, and mesh modeling. For instance, Dreamfusion by Google enables text-to-3D conversion using 2D diffusion techniques [46]. Dreamfields, a zero-shot text-guided object generation method, allows the generation of objects based on textual descriptions [24]. Shapenet, a comprehensive 3D model repository, provides rich information for 3D modeling tasks [8]. Advanced techniques like 3D GEN employ triplane latent diffusion for textured mesh generation [19], while Magic3D, developed by Nvidia, offers high-resolution text-to-3D content creation capabilities [29]. Point-e, an AI system by OpenAI, generates 3D point clouds from complex prompts [42]. RODIN, developed by Microsoft, is a generative model for sculpting 3D digital avatars using diffusion techniques [57]. MeshDiffusion focuses on score-based generative 3D mesh modeling [31], and Text2Mesh enables text-driven neural stylization for meshes [38]. Techniques such as zero-1-to-3 allow one image to be converted into a 3D object without any training data [30]. Additionally, EG-3D employs efficient geometry-aware generative adversarial networks for 3D modeling [7], and ICON/ECON enables implicit clothed human modeling from normals [59, 60]. These works highlight the diverse range of AI techniques and models that have been developed for various aspects of 3D scene creation. By incorporating these AI advancements, XR creators can unlock new possibilities and streamline their workflows, ultimately enhancing the immersive experiences delivered to users.

Besides assisting 3D content creation, AI can also be used to optimize and streamline various aspects of XR development, such as reducing the number of iterations required for 3D modeling and animation, and improving the efficiency of image processing and rendering tasks. These benefits can lead to significant time and cost savings, while also improving the quality of the final XR experience. Overall, AI can efficiently facilitate expanding and extending the creative possibilities of XR. By enabling more efficient and effective 3D content creation, AI can also help to drive the adoption and growth of XR technologies and unlock new opportunities for immersive and engaging experiences.

Figure 3: The application of AI-generated background environments in real-world filming and live testing.

4.2 Approaches for AI-assisted 3D Scene Creation in XR

To address the challenges of creating 3D scenes in XR, several approaches have been developed based on updated technologies, along with the use of software tools and platforms that incorporate AI technologies. These approaches leverage AI-based techniques to generate 3D models, textures, and lighting that are consistent with real-world scenes. Specifically, we summarize them into the following three categories and implemented the first one of them (Figure 3).

One of the key approaches to creating 3D scenes by AI is the use of High Dynamic Range (HDR) images. HDR images capture a wide range of lighting information, from bright highlights to dark shadows, which is essential for creating realistic lighting in 3D scenes. By using AI-based techniques, HDR images can be converted into fake 3D models, which can be further processed to generate 3D scenes. Tools like Skybox Lab [27] and Lumiere 3D [48] can be used for this purpose. However, this approach has limited perspectives. One potential solution to this limitation is to use multiple HDR images taken from different viewpoints to create a more comprehensive 3D model. This approach can be supported by 3D modeling software with AI plugins, such as Blender and Maya, which offer AI-assisted capabilities for tasks such as modeling, texturing, and rigging.

The second approach is based on 2D and involves the use of generative models such as GANs and Variational autoencoders (VAEs) [26] to generate 3D models and textures. These models can learn the underlying patterns and features of a given dataset of 3D models and textures, and then generate new ones that follow the same distribution. This allows for the rapid creation of a large number of unique models and textures with varying styles and features. For instance, Pix2Vertex [53] can generate 3D models from 2D images using a GAN. In addition to generating 3D models from existing assets, there are also methods that use existing 2D images or videos as a basis for 3D scene generation. This approach can benefit from AI-powered plugins in 3D modeling software, such as Blender's "Auto Eye" [13] addon for generating realistic eye textures and the "Graswald" [15] addon for realistic grass and vegetation.

Moreover, the third approach relies on constructing models using AI-based techniques. Apart from using simplified objects or predetermined shapes, these techniques enable the creation of intricate and lifelike 3D models, such as point cloud generation, which is capable of producing models of natural landscapes such as trees, rocks, and mountains. These models can subsequently be incorporated into larger 3D scenes, resulting in more immersive and realistic experiences. AI-based rendering tools, such as NVIDIA's Omniverse [44] and StyleProp [21], can be utilized to produce realistic and high-quality images and videos by using AI algorithms to simulate light and materials. These rendering software platforms often include AI-assisted capabilities and can be integrated into the workflow of 3D scene creation in XR.

One area where AI demonstrates its potential in the broader scope of 3D content creation is the exploration of fractals and procedural generation. Fractals, which exhibit intricate and self-similar patterns, have been utilized in various creative fields. In the context of 3D scene creation, artists can leverage tools like Blender to incorporate fractals into their designs, enabling the generation of visually captivating and mathematically-inspired scenes [34]. Moreover, AI-powered tools such as StyleProp [21] offer real-time example-based stylization of 3D models, providing artists with the means to infuse unique and artistic styles into their creations. Furthermore, advancements in audio recording and editing, such as Adobe's AI-powered audio processing [2], enhance the immersive experience of XR by seamlessly integrating realistic and high-quality audio elements. NVIDIA Canvas [43] and ControlNet [64] provide innovative solutions for stable diffusion and creative AI-based generation, respectively. Midjourney [40] and Reality OBJ to USDZ Converter [45] are tools that facilitate the transformation of assets into XR-compatible formats. Software like Adobe After Effects [1] and DaVinci Resolve [5] empower artists with professional editing capabilities, color grading, and audio post-production. RunwayML [52] serves as a platform to advance creativity with AI, offering a range of AI models and tools. References to image plane [49] and Trippy Everything AI music generation [12] demonstrate the fusion of AI and creative expression. Synthesis AI Labs [3] and ZBrush [36] provide cutting-edge AI-based solutions for tasks like text-to-3D conversion and digital sculpting. Monster Mash [18] and Spline AI [54] offer sketch-based modeling and animation tools as well as AI-powered object, animation, and texture generation. Additionally, advancements in audio-based AI models, such as Audiogpt [23], enable understanding and generation of speech, music, sound, and talking heads. Luma AI Capture [28] introduces lifelike 3D models and game art creation. Haque et al. [20] propose Instruct-NeRF2NeRF for editing 3D scenes with instructions. Finally, Mubert [58] offers AI-powered music generation for immersive XR experiences.

In summary, the creation of 3D scenes in XR involves the use of various AI-based approaches, supported by software tools and platforms. By leveraging HDR images, generative models, and AI-based techniques for model construction, artists and designers can create realistic and immersive 3D scenes. Additionally, the integration of AI plugins into 3D modeling software and the utilization of AI-based rendering software enhance the efficiency and quality of the creation process. XR development platforms with AI tools also provide a comprehensive solution by combining XR development capabilities with AI technologies. These advancements in AI-assisted 3D scene creation are expanding the possibilities of extended reality and expanded cinema experiences.

4.3 Case Study: Implementing AI-Generated HDR Images for Shooting

Inspired by the "Generate 3D Sets for your Short Films" project by YouTuber Mickmumpitz [39], we attempted to use text-to-image generation AI to create scenes and combine them with LED walls (in contrast to his use of green screens). As we aimed to film movies using visual production techniques, we were particularly interested in whether AI-generated scenes could be effectively utilized for real-life shooting scenarios in conjunction with live lighting. We conducted a rapid experiment to validate the feasibility of this.

According to the first approach mentioned in the paper, we want to demonstrate how AI-generated HDR images that can be leveraged to use as fake 3D scenes for visual production. Our initial experiment used a simple setup with an LED wall as the virtual background, a DSLR camera, and an AI-generated HDR image made by Skybox Lab. We used the latest scratch feature from Skybox Lab to create a basic structure for a 360 space and then used the SciFi mode to generate a surreal scene. In a very short time, we could project this scene onto our LED wall as a virtual background for filming. However, based on feedback, our initial experiment has some limitations and ways we can improve the methodology.

First, the conversion of the HDR image to a 3D model was done offline, making the experiment non real-time and the physical camera setup has not yet been aligned with the virtual camera in the software. This means we need to adjust the 3D perspectives manually in post-production on the computer to get the right perspective. So for future experiments, we aim to align them with real-time software and media servers used for virtual production to achieve synchronization. This will enable dynamic changes in perspective and movement within the scene. Once synchronized, the HDRI can effectively function as a real-time virtual scene but perhaps offers a limited range of perspectives due to calibration and synchronization challenges.

During our initial setup, we noted some key learnings, such as the ideal aspect ratio, resolution, the distance between the camera and LED wall, and camera settings. We also recorded the lighting conditions when using the AI-generated HDRI as a virtual scene. Here are some references to what we learned. For our nodal pan shot, from its center, we have room to pan left or right within a nearly 120 degree section of the screen avoiding moire. The aspect ratio we used was 16:9, and the resolution used was 1920x1080 pixels. The distance between the camera and the LED wall was 4 meters. The camera settings used an F-stop of f/4.5, ISO 800, and 50 fps. The ambient light level during the shooting was 800 lux, with indirect diffuse lighting. And we use additional lighting to complement the control of the front light. During this process, it is important to ensure that the logic and contrast of the front light are consistent with the content of the HDRI. Although our experiment was simple, we referred to the USC white paper [9], which discussed the best conditions to use LED as a virtual background and the lighting conditions.

The overall effect was quite impressive, with a focus on the characters' faces and a blurred background due to the use of a distant and close-up shot. In this case, the AI-generated scene was able to replace non-AI images. However, further research is needed to assess the effectiveness of AI-generated scenes in highly realistic environments or scenes that require highlighting of specific background details.

The main purpose of this study is to showcase the first step in AI to assist the potential of HDRI technology for expanding and extending cinematic experiences within virtual production workflows. These learnings will inform optimized parameters for our subsequent experiments.

5 FOR METAVERSE AND A MIGHT NEW FRONTIER

Puppetry and performance have creatively explored the intersection of the physical and virtual worlds for decades. And the shift from traditional 2D cinematic storytelling to 3D has been driven by technological advancements that made 3D cinematography possible [62]. But now, emerging technologies like virtual production, AI, and extended reality are increasingly blurring physical and virtual realities by integrating human participants with virtual characters and environments. Real-time servers converge multiple inputs to place high-resolution humans within virtual environments for interaction with virtual assets viewed. And improved LED and in-camera visual effects now enable new artistic expressions of real-time immersive and interactive effects.

These developments have also paved the way for expanded cinema, extended reality, and the emerging metaverse, which represent early strides toward realizing a future where vast virtual worlds seamlessly blend with the physical realm. These art forms could help redefine art expressions in a world that increasingly blends the virtual and the physical. Today, we focus on expanding and extending virtual experiences, like using Disguise Metaverse Labs [10] to create the next generation of immersive experiences through LED stages. And with the assistance of AI to push the limits of virtual experience further, generating hyper-realistic worlds, interactive characters, and deeply personalized realities. We believe they point to a future of seamless hybrid experience — one that blurs the lines of technology, media, and humanity itself. As these technologies continue to evolve, expanded cinema and other art forms will likely continue unfolding in new ways, shaping what it means to be human in the digital age.

REFERENCES

[1] Adobe. 2023. After Effects. https://www.adobe.com/products/aftereffects.html
[2] Adobe. 2023. Podcast. https://podcast.adobe.com/
[3] Synthesis AI. 2023. SynthesisAI. https://synthesis.ai/labs/
[4] Timothy Barker. 2012. Images and eventfulness: expanded cinema and experimental research at the University of New South Wales. *Studies in Australasian Cinema* 6, 2 (2012), 111–123.
[5] Blackmagicdesign. 2023. DaVinci Resolve. https://www.blackmagicdesign.com/products/davinciresolve
[6] Balthasar Blülle, Stéphane Altazin, Bérengère Frouin, Lidia Stepanova, Sandra Jenatsch, and Beat Ruhstaller. 2019. 37.4: Light conversion and scattering properties of QD films for display applications: Angle-resolved optical spectroscopy and numerical simulation. In *SID Symposium Digest of Technical Papers*, Vol. 50. Wiley Online Library, 407–410.
[7] Eric R Chan, Connor Z Lin, Matthew A Chan, Koki Nagano, Boxiao Pan, Shalini De Mello, Orazio Gallo, Leonidas J Guibas, Jonathan Tremblay, Sameh Khamis, et al. 2022. Efficient geometry-aware 3D generative adversarial networks. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*. 16123–16133.
[8] Angel X Chang, Thomas Funkhouser, Leonidas Guibas, Pat Hanrahan, Qixing Huang, Zimo Li, Silvio Savarese, Manolis Savva, Shuran Song, Hao Su, et al.

2015. Shapenet: An information-rich 3d model repository. *arXiv preprint arXiv:1512.03012* (2015).
[9] Electronic Theatre Controls. 2021. Ripple Effect White Paper. https://www.etcenter.org/wp-content/uploads/2021/03/Ripple-Effect-White-Paper-ETCUSC-March-2021.pdf
[10] Disguise. 2023. Discover the Metaverse. https://www.disguise.one/en/discover-the-metaverse/
[11] Disguise. 2023. Film and TV Drama XR Workshop: Five Ways XR is Revolutionising Film and TV. https://www.disguise.one/en/insights/blog/film-and-tv-drama-xr-workshop-five-ways-xr-is-revolutionising-film-and-tv/
[12] Trippy Everything. 2023. Trippy Everything. https://open.spotify.com/artist/0c8i3vyHL9XO5jdbIfNnig?si=DIGIC9q4R1adKx3YGfW_mA&nd=1
[13] Lucas Falcao. 2023. Auto eye. https://blendermarket.com/products/auto-eye
[14] Nametag Films. 2020. XR in Film Production. https://www.nametagfilms.com/xr-in-film-production
[15] Graswald GmbH. 2023. Graswald. https://www.graswald3d.com/
[16] Carmen Guiralt Gomar. 2014. El Two-Color Kodachrome y su inserción en The Llight in the Dark (Clarence Brown, 1922): primera y única exhibición pública del proceso experimental de color en un largometraje comercial. *Fotocinema: revista científica de cine y fotografía* 9 (2014), 82–116.
[17] Ian Goodfellow, Jean Pouget-Abadie, Mehdi Mirza, Bing Xu, David Warde-Farley, Sherjil Ozair, Aaron Courville, and Yoshua Bengio. 2020. Generative adversarial networks. *Commun. ACM* 63, 11 (2020), 139–144.
[18] Google. 2023. Monster Mash. https://monstermash.zone/
[19] Anchit Gupta, Wenhan Xiong, Yixin Nie, Ian Jones, and Barlas Oğuz. 2023. 3dgen: Triplane latent diffusion for textured mesh generation. *arXiv preprint arXiv:2303.05371* (2023).
[20] Ayaan Haque, Matthew Tancik, Alexei A Efros, Aleksander Holynski, and Angjoo Kanazawa. 2023. Instruct-nerf2nerf: Editing 3d scenes with instructions. *arXiv preprint arXiv:2303.12789* (2023).
[21] Filip Hauptfleisch, Ondrej Texler, Aneta Texler, Jaroslav Křivánek, and Daniel Sỳkora. 2020. StyleProp: Real-time Example-based Stylization of 3D Models. In *Computer Graphics Forum*, Vol. 39. Wiley Online Library, 575–586.
[22] Jonathan Ho, Ajay Jain, and Pieter Abbeel. 2020. Denoising diffusion probabilistic models. *Advances in Neural Information Processing Systems* 33 (2020), 6840–6851.
[23] Rongjie Huang, Mingze Li, Dongchao Yang, Jiatong Shi, Xuankai Chang, Zhenhui Ye, Yuning Wu, Zhiqing Hong, Jiawei Huang, Jinglin Liu, et al. 2023. Audiogpt: Understanding and generating speech, music, sound, and talking head. *arXiv preprint arXiv:2304.12995* (2023).
[24] Ajay Jain, Ben Mildenhall, Jonathan T Barron, Pieter Abbeel, and Ben Poole. 2022. Zero-shot text-guided object generation with dream fields. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*. 867–876.
[25] Marsha Kinder. 1999. Doors to the Labyrinth: Designing Interactive Frictions with Nina Menkes, Pat O'Neill, and John Rechy. *Style* 33, 2 (1999), 232–244.
[26] Diederik P Kingma and Max Welling. 2013. Auto-encoding variational bayes. *arXiv preprint arXiv:1312.6114* (2013).
[27] Blockade Labs. [n. d.]. Skybox AI. https://skybox.blockadelabs.com/
[28] Luma Labs. 2023. Luma AI Capture. https://lumalabs.ai/
[29] Chen-Hsuan Lin, Jun Gao, Luming Tang, Towaki Takikawa, Xiaohui Zeng, Xun Huang, Karsten Kreis, Sanja Fidler, Ming-Yu Liu, and Tsung-Yi Lin. 2023. Magic3d: High-resolution text-to-3d content creation. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*. 300–309.
[30] Ruoshi Liu, Rundi Wu, Basile Van Hoorick, Pavel Tokmakov, Sergey Zakharov, and Carl Vondrick. 2023. Zero-1-to-3: Zero-shot one image to 3d object. *arXiv preprint arXiv:2303.11328* (2023).
[31] Zhen Liu, Yao Feng, Michael J Black, Derek Nowrouzezahrai, Liam Paull, and Weiyang Liu. 2023. Meshdiffusion: Score-based generative 3d mesh modeling. *arXiv preprint arXiv:2303.08133* (2023).
[32] Jeremy Lovell. July 2003. Peter Tscherkassky: The Austrian Avant-Garde. http://www.sensesofcinema.com/2003/peter-tscherkassky-the-austrian-avant-garde/expanded_cinema/
[33] Batta Mahesh. 2020. Machine learning algorithms-a review. *International Journal of Science and Research (IJSR).[Internet]* 9, 1 (2020), 381–386.
[34] Jonas Mangelschots. 2023. Fractal Generator. https://blendermarket.com/products/fractal-generator
[35] B Marr. 2021. The fascinating history and evolution of extended reality (xr)–covering ar, vr and mr.".
[36] Maxon. 2023. ZBrush. https://www.maxon.net/en/zbrush
[37] MELO. 2023. Maurice Benayoun's Tunnel Under The Atlantic. https://oss.adm.ntu.edu.sg/melo0006/maurice-benayouns-tunnel-under-the-atlantic/
[38] Oscar Michel, Roi Bar-On, Richard Liu, Sagie Benaim, and Rana Hanocka. 2022. Text2mesh: Text-driven neural stylization for meshes. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*. 13492–13502.
[39] YouTuber Mickmumpitz. 2023. Generate 3D Sets for your Short Films! https://www.youtube.com/watch?v=t-8I7EkIL8c&t=535s
[40] Midjourney. 2023. Midjourney. https://www.midjourney.com/
[41] Ben Mildenhall, Pratul P Srinivasan, Matthew Tancik, Jonathan T Barron, Ravi Ramamoorthi, and Ren Ng. 2021. Nerf: Representing scenes as neural radiance fields for view synthesis. *Commun. ACM* 65, 1 (2021), 99–106.
[42] Alex Nichol, Heewoo Jun, Prafulla Dhariwal, Pamela Mishkin, and Mark Chen. 2022. Point-e: A system for generating 3d point clouds from complex prompts. *arXiv preprint arXiv:2212.08751* (2022).
[43] NVIDIA. 2023. NVIDIA Canvas. https://www.nvidia.com/de-de/studio/canvas/
[44] NVIDIA. 2023. Omniverse. https://www.nvidia.com/en-us/omniverse/
[45] Ariel Ortiz. 2023. Reality OBJ to USDZ Converter. https://apps.apple.com/de/app/reality-obj-usdz-converter
[46] Ben Poole, Ajay Jain, Jonathan T Barron, and Ben Mildenhall. 2022. Dreamfusion: Text-to-3d using 2d diffusion. *arXiv preprint arXiv:2209.14988* (2022).
[47] Samira Pouyanfar, Saad Sadiq, Yilin Yan, Haiman Tian, Yudong Tao, Maria Presa Reyes, Mei-Ling Shyu, Shu-Ching Chen, and Sundaraja S Iyengar. 2018. A survey on deep learning: Algorithms, techniques, and applications. *ACM Computing Surveys (CSUR)* 51, 5 (2018), 1–36.
[48] Jacquard Products. 2023. Lumiere 3D. https://www.lumiere3d.ai/
[49] Pullusb. 2023. Reference to image plane. https://github.com/Pullusb/reference_to_image_plane
[50] Catherine Rogers. 2014. *Film Outside Cinema.* Ph. D. Dissertation. Royal College of Art (United Kingdom).
[51] Julian A Ross. 2014. *Beyond the frame: intermedia and expanded cinema in 1960-1970s Japan.* Ph. D. Dissertation. University of Leeds.
[52] Runway. 2023. Runway.ml. https://runwayml.com/
[53] Matan Sela, Elad Richardson, and Ron Kimmel. 2017. Unrestricted facial geometry reconstruction using image-to-image translation. In *Proceedings of the IEEE International Conference on Computer Vision*. 1576–1585.
[54] Spline. 2023. Spline AI. https://spline.design/ai
[55] Stan VanDerBeek. 1966. "Culture: Intercom" and expanded cinema: A proposal and Manifesto. *Tulane Drama Review* 11, 1 (1966), 38–48.
[56] Jonathan Walley. 2020. *Cinema expanded: avant-garde film in the age of intermedia*. Oxford University Press.
[57] Tengfei Wang, Bo Zhang, Ting Zhang, Shuyang Gu, Jianmin Bao, Tadas Baltrusaitis, Jingjing Shen, Dong Chen, Fang Wen, Qifeng Chen, et al. 2023. Rodin: A generative model for sculpting 3d digital avatars using diffusion. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*. 4563–4573.
[58] Shangda Wu and Maosong Sun. 2022. Exploring the efficacy of pre-trained checkpoints in text-to-music generation task. *arXiv preprint arXiv:2211.11216* (2022).
[59] Yuliang Xiu, Jinlong Yang, Xu Cao, Dimitrios Tzionas, and Michael J Black. 2023. ECON: Explicit Clothed humans Optimized via Normal integration. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*. 512–523.
[60] Yuliang Xiu, Jinlong Yang, Dimitrios Tzionas, and Michael J Black. 2022. Icon: Implicit clothed humans obtained from normals. In *2022 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*. IEEE, 13286–13296.
[61] David Kei-Man Yip. 2020. The Invisible Art of Storytelling and Media Production. In *Advances in Creativity, Innovation, Entrepreneurship and Communication of Design: Proceedings of the AHFE 2020 Virtual Conferences on Creativity, Innovation and Entrepreneurship, and Human Factors in Communication of Design, July 16-20, 2020, USA*. Springer, 262–266.
[62] David Kei-Man Yip. 2022. Disruptive Innovations in Cinematic Storytelling from 2D to 3D. *Human Factors in Communication of Design* 49 (2022), 59.
[63] Gene Youngblood. 1970. Expanded Cinema. New York: P. Dutton & Co. *Inc. Search in* (1970).
[64] Lvmin Zhang and Maneesh Agrawala. 2023. Adding conditional control to text-to-image diffusion models. *arXiv preprint arXiv:2302.05543* (2023).

Received May 2023; accepted July 2023